Population dynamics of species-rich ecosystems: the mixture of matrix population models approach

Mortier, Frédéric; Rossi, Vivien; Guillot, Gilles; Gourlet-Fleury, Sylvie; Picard, Nicolas

Published in:
Methods in Ecology and Evolution

Link to article, DOI:
10.1111/2041-210x.12019

Publication date:
2013

Document Version
Early version, also known as pre-print

Population dynamics of species-rich ecosystems: the mixture of matrix population models approach

Frédéric Mortier[1*], Vivien Rossi[2], Gilles Guillot[3],

Sylvie Gourlet-Fleury[1] and Nicolas Picard[1]

[1]CIRAD, UPR Bsef, 34398 Montpellier, France

[2]CIRAD, UMR Ecofog, 97387 Kourou, Guyane, France

[3]Statistics Section IMM, Technical University of Denmark, Copenhagen, Denmark

[*]Correspondence author. E-mail: frederic.mortier@cirad.fr. Fax: (+33) 4 67 59 39 09

Running title: Mixture of matrix models

Word count: 6549

Summary

1. Matrix population models are widely used to predict population dynamics but, when applied to species rich ecosystems with many rare species, the small population sample sizes hinder a good fit of species-specific models. This issue can be overcome by assigning species to groups to increase the size of the calibration data sets. However, the species classification is often disconnected from the matrix modelling and from the estimation of matrix parameters, thus bringing species groups that may not be optimal with respect to the predicted community dynamics.

2. We proposed here a method that jointly classified species into groups and fit the matrix models in an integrated way. The model was a special case of mixture with unknown number of components and was cast in a Bayesian framework. An MCMC algorithm was developed to infer the unknown parameters: the number of groups, the group of each species and the dynamics parameters.

3. We applied the method to simulated data and showed that the algorithm efficiently recovered the model parameters.

4. When applied to a tree data set from a tropical rain forest in French Guiana, the mixture matrix model classified tree species into well differentiated groups with clear ecological interpretations. It also accurately predicted the forest dynamics over the sixteen-year observation period.

5. Our model and algorithm can straightforwardly be adapted to any type of matrix model, using the life cycle diagram. It can be used as an unsupervised classification technique to group species with similar population dynamics.

Key-words: Bayesian, clustering, mixture models, reversible jump Markov chain Monte

Carlo, tropical rainforests, species rich ecosystems, population dynamics.

Introduction

The conservation of animal and plant species and their biological control require models to understand and predict population dynamics (Fieberg & Ellner 2001; Buongiorno & Gilless 2003; Demyanov, Wood & Kedwards 2006). Among population dynamics models, projection matrix models have been widely used to investigate the dynamics of age-, stage- or size-structured populations (Caswell 2001; Stott, Townley, Carslake & Hodgson 2010). They provide a simple way of integrating vital rate information such as recruitment, birth, growth or ageing, and mortality (Crone, Menges, Ellis, Bell, Bierzychudek, Ehrlén, Kaye, Knight, Lesica, Morris, Oostermeijer, Quintana-Ascencio, Stanley, Ticktin, Valverde & Williams 2011). Matrix models have been used to model population demography in the context of species invasion (Hooten, Wikle, Dorazio & Royle 2007; Sebert-Cuvillier, Paccaut, Chabrerie, Endels, Goubet & Decocq 2007), species extinction or conservation of endangered species (Cropper & Loudermilk 2006), and the sustainable management of exploited species (Hauser, Cooch & Lebreton 2006). Recent improvements in matrix models targeted the estimation of demographic parameters, in particular for animal populations using capture-recapture methods (Besbeas, Freeman, Morgan & Catchpole 2002).

In species-rich ecosystems like tropical rain forests, tropical marine fish or coral reefs, high diversity implies that the number of individuals for most species is limited. The small sample size hinders a good fit of species-specific dynamics models, including matrix pop-

ulation models. To address this problem, modellers usually cluster species into groups. A variety of methods has been used to group species, favouring either ecological interpretation or the accuracy of predictions. Groups of species can be derived from functional groups (Steneck & Dethier 1994), ecomorphology (Bellwood & Wainwright 2001) or ecological subjective strategy (Swaine & Whitmore 1988; Favrichon 1994; Gitay & Noble 1997). None of these methods is a statistical method, thus not ensuring that the within-group similarity is maximum, or that the number of groups is optimal. Gourlet-Fleury, Cornu, Jésel, Dessard, Jourget, Blanc & Picard (2005) described two other strategies applied in tropical rain forests: the ecological data-driven strategy (Phillips, Yasman, Brash & van Gardingen 2002) and the dynamic process strategy, in which "process" refers to the components of forest dynamics (recruitment, growth or mortality) (Gourlet-Fleury & Houllier 2000; Picard, Mortier, Rossi & Gourlet-Fleury 2010). These strategies rely on statistical unsupervised classification methods, such as hierarchical cluster analysis, to group species with similar traits. Moreover, species classification is most often disconnected from the matrix modelling and from the estimation of the matrix parameters, thus bringing species groups that may not be optimal with respect to the predicted community dynamics.

The use of mixture models has recently been proposed to model the presence/absence of species (Dunstan, Foster & Darnell 2011), the species richness in a species assemblage (Mao, Colwell & Chang 2005), or the heterogeneity of capture and survival probabilities in natural populations (Pledger, Pollock & Norris 2010). Mixture models are based on the assumption that observation data arise from several unobserved groups (McLachlan & Peel 2000). A model is associated to each group. Each observation contributes to the fitting of the

model for a given group with a weight that represents its probability to belong to this group. These weights can eventually be used to classify observations among groups. Thus, mixture modelling simultaneously fits models and classifies observations, and the clustering step is closely linked to the calibration step. This favours the similarity of species response within groups rather than the similarity of species traits (Dunstan, Foster & Darnell 2011).

Mixture modelling has mainly been developed for observations with a normal distribution (e.g. mixture regressions). This study aims at extending mixture modelling to matrix population models. The mixture of matrix population models will simultaneously solve two issues: fit matrix models for species-rich ecosystem with many rare species, and classify species into groups. As proposed in population genetics (Pritchard, Stephens & Donnelly 2000; Corander, Waldmann & Sillanpaa 2003; Guillot, Estoup, Mortier & Cosson 2005), the strategy consists in a probabilistic model-based clustering method expressed in terms of matrix population mixture models with an unknown number of components (Richardson & Green 1997; Dunson 2000; Marin, Mengersen & Robert 2005). The number of groups and the parameters of the matrix population models associated with each group are the unknown quantities. We propose to use a Bayesian framework to infer these unknown quantities. The Bayesian framework approach has several advantages over frequentist methods. First, it enables us to obtain the exact posterior distribution for population sizes, whereas classical maximum likelihood methods provide asymptotic confident intervals. Secondly, with the use of prior distributions, strong biological or ecological knowledge can be included in the model.

The mixture of matrix models is defined in the next section. An inference method is then shortly presented, and extensively tested using simulated data. The mixture matrix model was

finally applied to a tree data set from the Paracou tropical rain forest in French Guiana. The tree species groups thus obtained had consistent ecological behaviours with contrasted functional traits, and compared favourably to other groups obtained by a standard classification technique.

Material and methods

Mixture of matrix population models

When fitting a base model to some observations, it is assumed that the set of observations is homogeneous, in the sense that all observations share a common distribution (e.g. the centred normal distribution for the residuals of the linear model). When dealing with an heterogeneous set of observations composed of K assumedly homogeneous subsets, finite mixture modelling is a relevant framework to extend this base model (McLachlan & Peel 2000). Finite mixture model assumes that the distribution of observations is a mixture of K base distributions, with mixing weights that represent the probability for an observation to belong to each of the homogeneous subsets. Conditionally on an observation belonging to a subset, the model identifies with the base model, while the distribution of the mixture includes the uncertainty on which subset an observation belongs to.

Finite mixture of matrix population models results from the application of the finite mixture framework to matrix population models. In matrix population models, individuals are classified into stage, size or age classes, and the population dynamics is described by transition rates among classes (Caswell 2001). At the individual level, these transitions can

be interpreted as the transitions of a Markov chain, which defines some distribution of the population-level numbers of individuals having passed between any two classes. Mixing K such distributions defines a mixture of K matrix population models. A specificity of the mixture of matrix models is that one observation corresponds to one population (more specifically, it is the vector of all numbers of individual transitions between classes), and the set of observations is the community-level set of populations. Hence, mixture of matrix models is relevant to model the dynamics of a community when assuming that its constituent species can be assigned to K homogeneous groups of species.

Hereafter, we detail the mathematical expression of the mixture of matrix models for a specific type of matrix population models, namely the Usher model. Nevertheless, this framework readily extends to any type of matrix models on the basis of individual transitions among classes.

Mixture of Usher matrix models

The Usher matrix model applies to size-structured populations (Usher 1966, 1969). It is based on the description of the change of the population by a vector, $\vec{N}_t$, of the numbers $N_{l,t}$ of individuals in L ordered size classes ($l = 1, \ldots, L$) at discrete time t. Let $N_t = \sum_{l=1}^{L} N_{l,t}$ be the total number of individuals at time t. Like any other matrix population model, the Usher model can be interpreted as the expectation of N_t independent Markov chains (Figure 1). The relationship between $\vec{N}_t$ and $\vec{N}_{t+1}$ is described by a $L \times L$ transition matrix U, called the Usher matrix:

$$\mathrm{E}[\vec{N}_{t+1}|\vec{N}_t] = U\,\mathrm{E}[\vec{N}_t] \tag{1}$$

where U is equal to:

$$U = \begin{pmatrix} p_1 + f & f & \dots & f \\ q_1 & p_2 & & 0 \\ & \ddots & \ddots & \\ 0 & & q_{L-1} & p_L \end{pmatrix} \tag{2}$$

p_l is the probability for an individual to stay in class l, q_l the probability to move up from class l to $l+1$ and f the average fecundity. q_l and p_l take values in $[0, 1]$, whereas f takes values in $\mathbf{R}^+$. The probability to die for an individual in class l is given by $m_l = 1 - p_l - q_l$. Let $\vec{d} = (d_1, \dots, d_L)$ be the class distribution of the population, such that d_l denotes the probability for a randomly chosen individual to belong to class l ($\sum_{l=1}^{L} d_l = 1$). Let $N_{l,l,t}$ denote the number of individuals staying in class l between $t-1$ and t, $N_{l,l+1,t}$ the number of individuals moving up from class l to $l+1$ between $t-1$ and t, and $N_{l,\dagger,t}$ the number of individuals dying in class l between $t-1$ and t. Let R_t be the number of recruits between $t-1$ and t, assumed to be a Poisson random variable with parameter fN_{t-1}. The vector of observations for the population is $\vec{N} = (N_{1,l,t}, \dots, N_{L,\dagger,t}, \vec{N}_{t-1}, R_t)$. The likelihood of the joined individual Markov transitions, and thus of the Usher matrix model, is:

$$\begin{aligned} \mathcal{L}(\vec{N}|\theta) \;=\; & \prod_{l=1}^{L-1} \mathcal{M}(N_{l,l,t}, N_{l,l+1,t}, N_{l,\dagger,t} | p_l, q_l, m_l, N_{l,t-1}) \\ & \times \mathcal{M}(N_{L,L,t}, N_{L,\dagger,t} | p_L, m_L, N_{L,t-1}) \\ & \times \mathcal{M}(N_{1,t-1}, \dots, N_{L,t-1} | d_1, \dots, d_L, N_{t-1}) \\ & \times \mathcal{P}(R_t | fN_{t-1}) \end{aligned} \tag{3}$$

where $\mathcal{M}$ denotes the multinomial distribution, $\mathcal{P}$ the Poisson distribution, and $\theta = (\vec{p}, \vec{q}, \vec{m}, f, \vec{d})$ is the vector of parameters with $\vec{p} = (p_1, \ldots, p_L)$, $\vec{q} = (q_1, \ldots, q_{L-1})$ and $\vec{m} = (m_1, \ldots, m_L)$. Eqn 1 is the deterministic version of the Usher projection model while eqn 3 accounts for the demographic stochasticity and is useful when the population size gets small (Caswell 2001).

Suppose now that the modelled population issues from K unobserved groups of species such that each group is modelled by a Usher projection matrix. Thus, there are K Usher matrices $U_1, \ldots, U_K$. Because the group to which the population belongs is not known a priori, one can define a random latent variable C that identifies the group of the species. For example, if the species belongs to the third group: $C = 3$. Conditionally on C, the prediction of the dynamics is given by eqn 1, with U being replaced by U_C. Accounting for the uncertainty on C brings:

$$\mathrm{E}[\vec{N}_{t+1}|\vec{N}_t] = \sum_{k=1}^{K} \pi_k \, U_k \mathrm{E}[\vec{N}_t] \tag{4}$$

where π_k is the posterior probability that C equals k. Eqn 4 defines the mixture of Usher matrix models, whose likelihood is:

$$\mathcal{L}(\vec{N}|\vec{\theta}, \vec{\pi}) = \sum_{k=1}^{K} \pi_k \, \mathcal{L}(\vec{N}|\theta_k) \tag{5}$$

where $\vec{\theta} = (\theta_1, \ldots, \theta_K)$ is the vector of all parameters associated with the K matrix models, $\vec{\pi} = (\pi_1, \ldots, \pi_K)$ is the vector of all posterior probabilities, and $\mathcal{L}(\vec{N}|\theta_k)$ is given by eqn 3. The species can be a posteriori classified by assigning it to the group g with the maximum posterior probability: $\pi_g = \max_k\{\pi_k\}$. Hence, the mixture of matrix models jointly defines

K matrix models (i.e. provides an estimate of $\vec{\theta}$) and classifies the species into K groups (i.e. provides an estimate of $\vec{\pi}$).

Mixture model inference

The parameters $\vec{\theta}$ and $\vec{\pi}$ of the mixture matrix model can be estimated in a frequentist context by maximizing the likelihood (5) of the mixture model. Inference can be achieved using an EM algorithm (McLachlan & Krishnan 2008). However, we here preferred to use Bayesian inference to have the opportunity to integrate biological knowledge into the model through the prior distribution of the parameters. Based on the direct acyclic graph of the mixture matrix model (Figure 2), a Markov chain Monte Carlo (MCMC) inference algorithm was implemented: a long sequence of parameter values was randomly drawn from the posterior distribution, and the parameter estimates were extracted from this sample by computing its mode or its means (Gilks, Richardson & Spiegelhalter 1996). Details on the Bayesian inference, including the choice of the priors, are given in Appendix A. Annotated R codes (R Core Team 2012) for the algorithm are available in the Supporting Information.

Fitting a finite mixture model also requires estimating the number K of groups. Classically, different mixture models with different number of groups are independently fitted, and an information criterion is finally used to select among these competing models (Biernacki, Celeux & Govaert 2000). A MCMC algorithm for a fixed K was developed with this aim in view. Alternatively, we also developed an inference algorithm that considered K as unknown and jointly estimated it with the other parameters. This involved using a reversible jump MCMC approach when the number of groups changed (Richardson & Green 1997). With

this latter approach, posterior probabilities for each value of K were obtained, thus enabling one to choose the most likely K while assessing the reliability of this choice.

Because the posterior distribution for the number K of groups may be sensitive to changes in the prior distribution for the parameters when using a reversible jump MCMC algorithm (Richardson & Green 1997), a sensitivity analysis to the priors was achieved. Details on the different priors that were tested are given in Appendix A.

Simulations

Data were simulated to assess the efficiency of the algorithm to correctly classify species into groups, according to different levels of differentiation between groups and different numbers of groups. Simulated data were composed of 100 species distributed across eight diameter classes. Numerical experiments tested the combinations of three factors: (*i*) the number of groups, that was equal to 1, 5 or 10 (3 modalities), and will be refered to as the true number of groups; (*ii*) the number of individuals per species, that was equal to 100 or 1000 (2 modalities); and (*iii*) hyper-priors for parameters ($\vec{d}$, $\vec{p}$, $\vec{q}$, $\vec{m}$, f), that took the values given in Table 1 (5 modalities).

The five different hyper-priors for the parameters corresponded to five levels of differentiation between groups. Indeed, the expectation of the diameter class or transition parameters was constant ($\mathrm{E}(d_l) = 1/8$ and $\mathrm{E}(p_l) = \mathrm{E}(q_l) = \mathrm{E}(m_l) = 1/3$ for all the hyper-priors in Table 1), but their variances decreased from 0.012 to 0.0015 for d_l and from 0.055 to 0.0079 for the transition parameters. As this variance corresponded to the between-group variance, the lower it was, the more similar the groups were. Let us note $\mathrm{Ldiff}_1, \ldots, \mathrm{Ldiff}_5$, the five de-

creasing differentiation levels of the hyper-parameters. When the number of groups was one, only the level Ldiff_1 was used for hyper-priors. In total, there were thus: $2\times1+2\times2\times5 = 22$ combinations of factors in the numerical experiments. For each combination, 50 replications were simulated. For each replication, the 100 species were randomly assigned to groups. This simulated classification was the reference to compare with the estimated classification and was referred as the 'true classification'. Then, for each group, the diameter class parameters, the transition parameters and the fecundity parameter were randomly drawn according to their hyper-prior distributions (Table 1). Finally, for each species, the prescribed number of individuals was drawn according to the law defined by eqn 3 using the parameters of the group to which the species belonged.

To assess the performance of the method, we compared the estimated number $\hat{K}$ of groups with the true number K used to simulate data sets, and we compared the estimated classification with the true classification using two set matching indices I_1 and I_2 (Meilă 2007). These indices are based on the $K \times \hat{K}$ contingency table $T = (T_{ij})$ with $i = 1, \ldots, K$ and $j = 1, \ldots, \hat{K}$ that cross-tabulates the species according to the true and the estimated classifications:

$$I_1 = \frac{1}{S}\sum_{i=1}^{K} \max\left\{T_{i1}, \ldots, T_{i\hat{K}}\right\} \quad \text{and} \quad I_2 = \frac{1}{S}\sum_{j=1}^{\hat{K}} \max\left\{T_{1j}, \ldots, T_{Kj}\right\}$$

These indices vary between $1/S$ and 1, and the higher they are, the better is the adequacy between the two classifications (Meilă 2007). They jointly reflect how groups collapsed and merged: $I_1 = 1$ and $I_2 = 1$ means that both classifications were identical; $I_1 = 1$ and $I_2 < 1$ means that the number of groups was underestimated and one or more groups were merged;

$I_1 < 1$ and $I_2 = 1$ means that the number of groups was overestimated and one or more groups were split; $I_1 < 1$ and $I_2 < 1$ means that several set operations are needed to move from one classification to the other.

Tropical forest data

Data on the tropical rain forest were collected at the Paracou experimental site ($5°18'$N, $52°53$'W), French Guiana. The site is located in a undisturbed *terra firme* forest under equatorial climate. Three 250 m $\times$ 250 m permanent sample plots (18.75 ha in total) have been established in 1984 and left as control of the undisturbed forest dynamics. All trees greater than 10 cm dbh (diameter at breast height) have been identified and georeferenced. Girth at breast height, standing deaths, treefalls and newly recruited trees greater than 10 cm dbh have been monitored either annually or every two years since 1984 (Gourlet-Fleury, Guehl & Laroussinie 2004). Because the Paracou forest is a mature undisturbed forest, the diameter distribution in those control plots could be considered at quasi-equilibrium. Two data sets were extracted from the Paracou database: one training data set to infer the mixture of Usher models, and one validation data set. A data set gave the species, the diameter class at year t and the diameter class at year $t + 2$ for n trees. Trees that died between years t and $t + 2$, and trees whose diameter overcame the inventory threshold of 10 cm between years t and $t + 2$ (recruited individuals) were included in the data set.

The training data set consisted of the data collected in 1993 and 1995 on the three control plots. One hundred and eighty one species were identified in these three control plots (Figure 3), illustrating both the high species richness, and the relative scarcity of most species

of the Guianan forest. The mean number of individuals per species was 64.54 (total on the three control plots of the training data set), with a minimum of 1 and a maximum of 980. The median number of individuals per species was 22, with a first quartile of 8 and a third quartile of 61.25. Although it could be possible to include species with few individuals into the analysis, we decided to leave out species with less than 20 individuals in the control plots in 1993. A preliminary analysis (not shown) evidenced that there was little difference between the classification based on all species and the classification restricted to species having at least 20 individuals: the algorithm took longer to converge in the former case, rare species were not well classified, and actually behaved like noise with respect to the estimation of groups. Moreover, from an ecological point of view, it does not make sense to assign species to groups when they are represented by few individuals. It is ecologically much more meaningful to a posteriori assign rare species to existing groups, using expert's knowledge on the species autecology. Hence, we reckon that rare species should rather be a posteriori assigned to existing groups. We were left with 93 species that included at least 20 trees monitored in the three control plots. This training data set contained 10,756 trees. The validation data set consisted of the data collected in 2009 on the same three control plots.

A classification of tree species into five groups was defined at Paracou by Favrichon (1994), using multivariate analysis and k-means clustering of species attributes (including size summary statistics, growth and recruitment). On the basis of these groups, Favrichon (1998) then fitted a Usher matrix model to predict forest dynamics. Hence, Favrichon's approach is illustrative of a two-step approach with a species classification that is disconnected from the matrix population model. We compared Favrichon's species classification with the

one obtained by the mixture matrix model using the likelihood (5) of the training data set. Because there were missing observations between 1995 and 2009, the same computation was intractable for the validation data set. Nevertheless, considering that the undisturbed forest was close to equilibrium, we also compared the likelihoods of the validation data set given the asymptotic diameter distributions according to the two classifications. For a given population with Usher transition matrix U (eqn 2), the asymptotic diameter distribution is the normalized eigenvector of U associated to its dominant eigenvalue (Caswell 2001).

Results

Recovery of simulated classifications

Simulation results were similar whether we used a uniform or a truncated Poisson distribution as a prior for K. Hence, only the results with the later prior (that was the default one) are reported here. For 1000 individuals per species, the estimated classification perfectly matched with the true simulated classification for all differentiation levels: I_1 and I_2 were always equal to one.

For 100 individuals per species, the results depended on the differentiation levels and on the number of groups (Table 2). When the true number of groups was one, the algorithm always found one group. For 5 groups, we correctly estimated the number of groups in 100, 100, 96, 76 and 52$\%$ of the cases for the 5 decreasing levels of differentiation respectively. When the number of groups was wrongly estimated, it was systematically underestimated: I_1 was very close to 1 and I_2 always remained lower than I_1. The classification method tended

to merge different species groups into one group, and to dispatch very few species of a given group into another group. The same results were found with stronger evidence in the case of 10 groups. At the fourth level of differentiation, the number of group was correctly estimated in about 80% of the cases, and more than 95% of the species were classified into the correct groups.

Tropical rain forest tree species classification

The 93 tree species at Paracou were classified using the mixture of matrix models, based on eight diameter classes ($\leq$ 15 cm, 15–20, 20–25, 25–30, 30–40, 40–50, 50–60, $\geq$ 60 cm). Based on 50 different chains, and 20,000 iterations after a burn-in of 10,000 iterations, five groups were obtained 48 times and six groups twice. Groups remained globally the same for all chains. We kept the chain with the highest log-likelihood. For this chain, the posterior probabilities for $K = 5$, 6, 7 or 35 groups were equal to 0.99, 5.3×10^{-3}, 9.3×10^{-4} and 6.7×10^{-5}, respectively.

The sensitivity analysis to the prior distributions showed that the estimate of K was fairly insensitive to the specification of the prior distributions for the parameters. For all priors except one, the algorithm found again five groups of species. The exception corresponded to $\alpha = \beta = 10$ for the priors of the transition and diameter class parameters, to be compared to $\alpha = \beta = 1$ for the default prior (Appendix A). In that case, K was estimated to three groups (with former groups 2 and 3 merged into a single one, and former groups 4 and 5 merged into a single one). Because α and β can be interpreted as pseudo-counts of individuals in diameter classes, large values of α and β tend to decrease the impact of observations

on the classification, in particular for the largest diameter class that have few observations. Hence, the sensitivity of K to α and β expresses the sensitivity of the species classification to differences between species in the largest diameter classes.

To help interpreting the five species groups, five demographic and biological attributes were computed for each group: growth rate, mortality rate, fecundity rate, upper bound for diameter, and turnover. Direct estimates of these attributes were computed from the training data set, and compared to the indirect estimates obtained from the estimated transition and diameter class parameters of the mixture matrix model (see the Supplementray Information for the estimates of all mixture matrix model parameters). The direct estimate of growth was the mean diameter increment between 1993 and 1995 of all trees that belonged to the group, while its indirect estimate was $\sum_{i=1}^{L-1} p_i d_i \delta_i$, where δ_i is the width of the ith diameter class. The direct estimate of the mortality was the ratio of the number of dead trees in the group between 1993 and 1995 over the number of trees in the group in 1993, while its indirect estimate was $\sum_{i=1}^{L} m_i d_i$. The direct estimate of the fecundity was the ratio of the number of recruited trees in the group between 1993 and 1995 over the number of trees in the group in 1993, while its indirect estimate was f. The direct estimate of the upper bound for diameter was the 95% quantile of diameters in 1995, while its indirect estimate was interpolated from $\vec{d}$ assuming that the diameter distribution was uniform within each class. Finally, the turnover was computed as half the sum of the mortality rate and of the fecundity rate. The direct and indirect estimates of these attributes were not expected to be strictly equal since they did not derive from the same estimators; yet, their values were quite close and evidenced the same differences between groups (Table 3).

Groups were labelled by decreasing order of growth (Table 3). The gradients of maximum size and turnover perfectly paralleled this gradient of growth, with the fastest growing group 1 having the greatest maximum size and the lowest turnover rate. Group 1 was composed of emergent mid-tolerant species, i.e. species that need to settle in the upper strata and sometimes above the forest canopy to complete their whole life-cycle. Group 2 was composed of a mix of shade-tolerant (mostly) and light-demanding (to a lesser extent) canopy species. Group 3 was composed of shade-tolerant species, with a mix of canopy (mostly) and understorey (to a lesser extent) species. As a consequence, its growth rate and maximum size were lower than for group 2, but higher than for group 4. The two small-sized groups 4 and 5 were composed of understorey shade-tolerant species, although group 4 also included a few pioneer species. As a consequence, the growth rate of group 4 was higher than that of group 5.

Because mixture of matrix models jointly classifies species and fits matrix models, we also compared the predicted and the observed number of individuals in each diameter class and each group in 2009, to check the validity of the matrix model. The mixture matrix population model correctly predicted both the number of trees 16 years later and their size distribution (Figure 4).

The log-likelihood of the training data set was -2722.7 for the Bayesian classification and -3351.7 for Favrichon's classification. The log-likelihood of the validation data set given the asymptotic diameter distriution was -2007.7 for the Bayesian classification and -2874.3 for Favrichon's classification. Hence, both criteria largely favoured the Bayesian classification to the detriment of Favrichon's classification.

Discussion

Mixture modelling can deal with matrix population models, and can jointly classify species and fitting matrix models. Mixture of matrix population models can be addressed in the frequentist or in the Bayesian context. The algorithm that we developed in the Bayesian context performed well on simulated data with known groups, even when the differentiation between groups was low. Classification was correctly predicted when between-group variances were higher than 0.0019 for diameter parameters ($\vec{d}_k$) and 0.010 for transition parameters ($\vec{p}_k, \vec{q}_k, \vec{m}_k$ and f_k), corresponding to the fourth level of differentiation (see Table 1). A specificity of the Bayesian method presented here is that it estimated the number K of groups together with the other parameters. This is quite original as mixture modelling generally operates conditionally on K, and then uses an information criterion to select K (Biernacki, Celeux & Govaert 2000). Moreover, the Bayesian approach allowed us to construct prior distributions taking into account ecological expert knowledge. For example, we assumed that the prior diameter distribution was a Dirichlet distribution where all parameters were equal to one meaning that the diameter distribution was uniform across diameter classes. Nevertheless, using the Bayesian paradigm, it is straightforward to change the prior distribution to model expert knowledge, assuming for example that diameter distribution is decreasing from the first to the last diameter class. Another example was the use of the prior distribution used to model recruitment based on expert knowledge

The method that we developed for the mixture of Usher matrix models could straightforwardly be adapted to other types of matrix projection models, such as Leslie or Lefkovitch

matrix models for age- and stage-structured populations, respectively. Starting from the life cycle representation of the matrix model (Figure 1), one simply has to translate the probabilities associated to each transition into a distribution law for an observation (eqn 3).

When applied to a tropical rainforest at Paracou, the mixture of Usher matrix models was able to jointly classify species and make reliable predictions. Predictions were better with the mixture model than with Favrichon's two-step approach, thus exemplifying that a classification disconnected from the matrix model may not be optimal to predict the community dynamics. The characteristics of the tree species groups formed at Paracou were consistent with known ecological behaviour (Lieberman, Lieberman, Hartshorn & Peralta 1985; Nascimento, Laurance, Condit, Laurance, D'Angelo & Andrade 2005; Delcamp, S., O. & E. 2008; Poorter, Wright, Paz, Ackerly, Condit, Ibarra-Manríquez, Harms, Martínez-Ramos, Mazer, Muller-Landau, Peña-Claros, Webb & Wright 2008): small-sized species (with the exception of pioneers) tend to grow slowly, to have high recruitment and mortality rates (i.e. high turnover rates), whereas large sized species that reach the forest canopy tend to grow rapidly and have low turnover rates. The mixture of Usher matrix models classified species according to both their growth rate and their maximum size (Picard, Köhler, Mortier & Gourlet-Fleury 2012). When plotting species along these two axes, species groups were clearly separated (Figure 5). Because these two axes can be used to order species along a continuum of ecological strategies (Turner 2001; Alder, Oavika, Sanchez, Silva, Van der Hout & Wright 2002), this means that the mixture of Usher matrix models was also able to classify species in a way that is consistent with their autecology.

The heterogeneity, in terms of light-requirement, found in groups 2 and 4 can be eas-

ily understood given the environmental conditions prevailing in the control plots. These plots are largely undisturbed, with only small gaps occurring at a rate of more or less 3 per year (Gourlet-Fleury, Guehl & Laroussinie 2004). Such conditions do not favour the growth of light-demanding species, nor the growth and survival of pioneer species. Because these species do not express their growth potential, they tended to be gathered with slower-growing species in groups 2 and 4. This, in addition to the fact that few pioneer species can survive in these plots, explains why no pioneer group was identified by our procedure while such a group usually is the first one to be isolated in a classification, due to its particular behaviour (Swaine & Whitmore 1988). Applying the mixture of matrix models to disturbed plots would have raised a different classification better accounting for the variety of potential specific behaviours.

In the Paracou example, the distribution of individuals across diameter classes in 1993 was taken into account in the mixture of matrix models: the likelihood (eqn 3) depended on the vector of parameters $\vec{d}$. This means that the shape of the initial diameter distribution influenced the outcome of the species classification. This made sense for the Paracou control plots because these plots were settled in undisturbed forest, whose state in 1993 could be considered as close to equilibrium. The vector $\vec{d}$ was thus representative of the equilibrium state of the forest. We checked indeed (results not shown here) that the asymptotic growth rate of the matrix models were close to one, and the associated eigenvectors close to $\vec{d}$. In other situations where the forest is far from equilibrium, it might not be advisable to account for the initial diameter distribution $\vec{d}$ in the species classification. Computing the conditional likelihood knowing $\vec{N}_t$ would enable to drop $\vec{d}$ from the expression of the likelihood (eqn 3).

390 Apart from this, the mixture of matrix models would be unchanged.

391 Appendix A Bayesian inference

Let S be the number of species in the calibration data set. Using the same notation as above with the additional superscript s, let $\vec{N}^s = (N^s_{1,l,t}, \ldots, N^s_{L,\dagger,t}, \vec{N}^s_{t-1}, R^s_t)$ be the vector of observations for species $s = 1, \ldots, S$ and let $\underline{N} = (\vec{N}^1, \ldots, \vec{N}^S)$ be the vector of observations for all species. Let $\vec{C} = (C_1, \ldots, C_S)$ be the latent vector that gives the group of each species. Considering K as unknown, the posterior probability π_k follows from the posterior density distribution of the mixture model:

$$\pi^{\underline{N}}_{\vec{C},\vec{\theta},K}(\vec{C}, \vec{\theta}, K|\underline{N}) \propto \prod_{s=1}^{S} \mathcal{L}(\vec{N}^s|\theta_{C_s})\ \pi^0_{\vec{C}|\vec{\theta},K}(\vec{C}|\vec{\theta}, K)\ \pi^0_{\vec{\theta}|K}(\vec{\theta}|K)\ \pi^0_K(K) \tag{6}$$

392 where $\mathcal{L}(\vec{N}^s|\theta_{C_s})$ is given by eqn 3, and $\pi^0_{\vec{C}|\vec{\theta},K}$, $\pi^0_{\vec{\theta}|K}$ and π^0_K are the prior densities associated
393 with the class latent random variables, the parameters of each matrix model and the number
394 of groups, respectively. For full Bayesian inference of the model, we set the followings priors
395 on the unknown quantities $\vec{C}$, $\vec{\theta}$ and K.

396 We assumed that the prior distribution for the number K was a Poisson distribution with
397 mean one, truncated to strictly positive values: $\pi^0_K(K) \equiv \mathcal{P}(1)\backslash\{0\}$. This prior distribution
398 was suggested by Nobile (2005) in order to be more parsimonious than under uniform distri-
399 bution. For the sensitivity analysis, a uniform distribution between one and S was also used
400 as a prior for K.

The parameters associated with the matrix population model for group k are $(\vec{p}_k, \vec{q}_k, \vec{m}_k)$, f_k and $\vec{d}_k$. The prior for the parameters $\vec{\theta}$ of the K matrix population models assumed that the

parameters of the different classes and groups were independent:

$$\pi^0_{\vec{\theta}|K}(\vec{\theta}|K) = \prod_{k=1}^{K} \left\{ \prod_{l=1}^{L-1} \pi^0_{p,q,m|l,k}(p_{lk}, q_{lk}, m_{lk}) \right\} \pi^0_{p,m|k}(p_{Lk}, m_{Lk})\ \pi^0_{\vec{d}|k}(\vec{d_k})\ \pi^0_{f|k}(f_k)$$

Because the Dirichlet distribution (denoted $\mathcal{D}$) is the conjugate prior of the multinomial distribution, we used the Dirichlet distribution as a prior for all transition parameters and all diameter class parameters: $\pi^0_{\vec{d}|k} \equiv \mathcal{D}(\alpha, \ldots, \alpha)$, $\pi^0_{p,q,m|l,k} \equiv \mathcal{D}(\beta, \beta, \beta)$ and $\pi^0_{p,m|k} \equiv \mathcal{D}(\beta, \beta)$, where α and β are hyper-parameters that can be interpreted as pseudo-counts of individuals. The default priors used $\alpha = \beta = 1$. For the sensibility analysis, we also tested $\alpha = \beta = 0.5$ that corresponds to the non-informative Jeffreys prior (Jeffreys 1946; Atwood 1996), and $\alpha = \beta = 10$. Because the gamma distribution (denoted $\mathcal{G}$) is the conjugate prior of the Poisson distribution, we used the gamma distribution as a prior for the fecundity parameter: $\pi^0_{f|k} \equiv \mathcal{G}(\gamma, \delta)$, where δ and γ are hyper-parameters. The default prior used $\gamma = 0.01$ and $\delta = 1$, which expresses the expert's knowledge that the recruitment rate in undisturbed natural rain forest is around 1%. For the sensitivity analysis, we also tested $\gamma = 0.5$ and $\delta = 1$, 10^{-1} or 10^{-10} (but the Jeffreys prior that corresponds to $\gamma = 0.5$ and $\delta = 0$ could not be used because it is improper).

The prior for the class vector $\vec{C}$ assumed that, given the number of groups, each species could equally and independently of the other species be in any group: $\pi^0_{\vec{C}|\vec{\theta},K}(\vec{C}|\vec{\theta}, K) = \prod_{s=1}^{S} \pi^0_{\vec{C}|K}(C_s|K)$ where $\pi^0_{\vec{C}|K}(C_s|K)$ is a uniform distribution on the number of groups: $\mathcal{U}(1, \ldots, K)$.

The inference of parameters was made through the investigation of the posterior distribution $\pi^{\underline{N}}_{\vec{C},\vec{\theta},K}(\vec{C}, \vec{\theta}, K|\underline{N})$ defined by eqn 6. As the number of groups was unknown, the

posterior distribution was not available in an analytic form. Hence, a specific Metropolis within Gibbs Monte Carlo Markov chain (MCMC) algorithm was developed. The algorithm consisted of three moves: increasing the number of groups (birth case); decreasing the number of groups (death case); keeping the same number of groups but potentially changing one species assignment (no jump case). In the first two cases, the number of parameters was not constant, so a reversible jump MCMC approach was used (Richardson & Green 1997), whereas in the third case, a Gibbs step could be used. All moves were equally distributed with probability 1/3.

In the following, we detail the proposal step for the three moves and the selection step for the birth and death cases.

1. Proposal step. Let us $|k|$ denote the number of species in group k, for $k = 1, \ldots, K$. Let us $K^\star$ denote the number of groups of the proposal and $\vec{C}^\star$ denote the latent class vector of the proposal.

   - No jump case: $K^\star = K$. The proposal $\vec{C}^\star = (C_1^\star, \ldots, C_S^\star)$ for the latent class vector is drawn in two steps:

     (a) randomly choose one species s among the groups that include two or more species;

     (b) new assignment $C_s^\star$ for species s is sampled from a multinomial distribution $\mathcal{M}(1;\, w_1, \ldots, w_K)$, whereas $C_t^\star = C_t$ for $t \neq s$. The coefficients w_k are equal to

$$w_k = \frac{\mathcal{L}(\vec{N}^s|\theta_k)}{\sum_{j=1}^{K} \mathcal{L}(\vec{N}^s|\theta_j)}$$

where $\mathcal{L}$ is given by (3).

- Birth case: $K^\star = K + 1$. The proposal for the latent class vector is obtained by splitting one group into two subgroups:

  (a) randomly choose one group k among the groups that include two or more species; this group will form two subgroups labelled k_1 and k_2;

  (b) choose the number $|k_1|$ of species that will compose group k_1 following a uniform distribution: $|k_1| \sim \mathcal{U}(1, \ldots, |k| - 1)$

  (c) sample $|k_1|$ species among the $|k|$ species in group k and allocate them to the first subgroup k_1. The others are allocated to the second subgroup k_2. Let D denote the resulting allocation vector of the $|k|$ species between k_1 and k_2.

  Let $\vec{C}^\star = (\vec{C}, k, |k_1|, D)$ denote the new classification that results from $\vec{C}$ through steps (a)–(c). Then, the conditional probability distribution of the new classification into $K + 1$ groups given the old one into K groups, $\pi^{\text{split}}_{\vec{C}^\star|\vec{C},K}$, is defined by:

$$\begin{aligned} \pi^{\text{split}}_{\vec{C}^\star|\vec{C},K}(\vec{C}^\star|\vec{C}, K) &= \Pr(\vec{C}^\star = (\vec{C}, k, |k_1|, D)|\vec{C}, K) \\ &= \frac{|k_1|!(|k| - |k_1|)!}{|k|!} \frac{1}{|k| - 1} \frac{1}{\sum_{i=1}^{K} \mathbb{1}_{|i|>1}} \frac{1}{2} \end{aligned}$$

- Death case: $K^\star = K - 1$. The proposal for the latent class vector is obtained by merging two groups into a single one: randomly choose two groups among K and merge them into one group. Let k_1 and k_2 be the two selected groups and let $\vec{C}^\star = (\vec{C}, k_1, k_2)$ be the new classification that results from $\vec{C}$ by merging k_1 and k_2. Then, the conditional probability distribution of the new classification into

$K-1$ groups given the old one into K groups, $\pi^{\text{merge}}_{\vec{C}^\star|\vec{C},K}$, is defined by:

$$\begin{aligned}\pi^{\text{merge}}_{\vec{C}^\star|\vec{C},K}(\vec{C}^\star|\vec{C},K) &= \Pr(\vec{C}^\star=(\vec{C},k_1,k_2)|\vec{C},K)\\ &= \frac{2!(K-2)!}{K!}\frac{1}{2}\end{aligned}$$

2. Selection step. Given $\vec{C}$ and K, the vector of new parameters $\vec{\theta}^\star=(\vec{p}^\star,\vec{q}^\star,\vec{m}^\star,f^\star,\vec{d}^\star)$ is sampled from its marginal posterior distribution $\pi^{\underline{N}}_{\vec{\theta}|\vec{C},K}(\vec{\theta}|\vec{C},K,\underline{N})$. This marginal posterior distribution (not given here to save space) is known in an analytical form since multinomial/Dirichlet and Poisson/gamma distributions are conjugate distributions (Robert & Casella 2005).

The following equations give the expression of the Metropolis-Hasting ratio in the death case, for example. Let the current number of groups be K, and the new state $K^\star$ be $K-1$. Let us assume that two groups k_1 and k_2 have been chosen and merged into a unique group k. Then,

$$\frac{\pi^{\text{split}}_{\vec{C}|\vec{C}^\star,K^\star}(\vec{C}|\vec{C}^\star,K^\star)}{\pi^{\text{merge}}_{\vec{C}^\star|\vec{C},K}(\vec{C}^\star|\vec{C},K)}=\frac{\binom{|k|}{|k_1|}\frac{1}{|k|-1}\frac{1}{\sum_{i=1}^{K}\mathbb{1}_{|i|>1}}}{\binom{K}{2}}$$

Moreover, $\frac{\pi^{\underline{N}}_{\vec{\theta}|\vec{C},K}(\vec{\theta}|\vec{C},K,\underline{N})}{\pi^{\underline{N}}_{\vec{\theta}|\vec{C},K}(\vec{\theta}^\star|\vec{C}^\star,K^\star,\underline{N})}$ is the ratio of marginal posterior distributions of $\vec{\theta}$ and is equal to

$$\frac{\pi^{\underline{N}_k}_{\theta}(\theta_k|\underline{N}_k)}{\pi^{\underline{N}_{k_1}}_{\theta}(\theta_{k_1}|\underline{N}_{k_1})\pi^{\underline{N}_{k_2}}_{\theta}(\theta_{k_2}|\underline{N}_{k_2})}$$

where $\underline{N}_k$ is the set of observations belonging to all species classified in group k. $\pi^{\underline{N}_k}_{\theta}(\theta|\underline{N}_k)$ is broken down as follows:

$$\pi^{\underline{N}_k}_{\theta}(\theta|\underline{N}_k)=\prod_{l}^{L}\pi^{\underline{N}_k}_{pqm|l,k}(p_l,q_l,m_l|\underline{N}_k)\;\pi^{\underline{N}_k}_{\vec{d}|k}(\vec{d}|\underline{N}_k)\;\pi^{\underline{N}_k}_{f|k}(f|\underline{N}_k)$$

where

$$\pi^{\frac{N_k}{pqm|l,k}} \equiv \mathcal{D}(1 + n_{llk}, 1 + n_{l(l+1)k}, 1 + n_{l\dagger k})$$

where n_{llk}, $n_{l(l+1)k}$ and $n_{l\dagger k}$ are the number of individuals in group k that respectively stay in class l, move from class l to $l + 1$ or die;

$$\pi^{\frac{N_k}{\vec{d}|k}} \equiv \mathcal{D}(1 + n_{lk}, \ldots, 1 + n_{Lk})$$

where n_{lk} is the number of individuals of group k in class l at initial time t; and finally,

$$\pi^{\frac{N_k}{f|k}} \equiv \mathcal{G}\left(0.01 + n_{01k}, \frac{1}{n_k + 1}\right)$$

where n_k is the total number of individuals in group k at initial time t and n_{01k} is the number of recruits in group k. Given this, the calculation of prior distribution as well as likelihood ratios is straightforward. As the matrix population model parameters are sampled from their posterior distributions, the canonical reversible transition function is the identity function. Hence, its Jacobian is equal to one and does not appear in the Metropolis-Hasting ratios.

References

Alder, D., Oavika, F., Sanchez, M., Silva, J., Van der Hout, P. & Wright, H. (2002) A comparison of species growth rates from four moist tropical forest regions using increment-size ordination. *International Forestry Review*, **4**(3), 196–205.

Atwood, C.L. (1996) Constrained noninformative priors in risk assessment. *Reliability Engineering and System Safety*, **53**(1), 37 – 46.

Bellwood, D. & Wainwright, P. (2001) Locomotion in labrid fishes: implications for habitat use and cross-shelf biogeography on the great barrier reef. *Coral Reefs*, **20**(2), 139–150.

Besbeas, P., Freeman, S.N., Morgan, B.J.T. & Catchpole, E.A. (2002) Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*, **58**(3), 540–547.

Biernacki, C., Celeux, G. & Govaert, G. (2000) Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(7), 719–725.

Buongiorno, J. & Gilless, J.K. (2003) *Decision Methods for Forest Resource Management*. Academic Press, Amsterdam, The Netherlands.

Caswell, H. (2001) *Matrix Population Models: Construction, Analysis and Interpretation*, 2nd edn. Sinauer, Sunderland, Massachusetts.

Corander, J., Waldmann, P. & Sillanpaa, M. (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367–374.

Crone, E.E., Menges, E.S., Ellis, M.M., Bell, T., Bierzychudek, P., Ehrlén, J., Kaye, T.N., Knight, T.M., Lesica, P., Morris, W.F., Oostermeijer, G., Quintana-Ascencio, P.F., Stanley, A., Ticktin, T., Valverde, T. & Williams, J. (2011) How do plant ecologists use matrix population models? *Ecology Letters*, **14**, 1–8.

Cropper, W. & Loudermilk, E. (2006) The interaction of seedling density dependence and

fire in a matrix population model of longleaf pine (*Pinus palustris*). *Ecological Modelling*, **198**, 487–494.

Delcamp, M., S., G.F., O., F. & E., G. (2008) Can functional classification of tropical trees predict population dynamics after disturbance? *Journal of Vegetation Science*, **19**(2), 209–220.

Demyanov, V., Wood, S.N. & Kedwards, T.J. (2006) Improving ecological impact assessment by statistical data synthesis using process-based models. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **55**(1), 41–62.

Dunson, D.B. (2000) Bayesian latent variable models for clustered mixed outcomes. *Journal of the Royal Statistical Society*, **62**(2), 335–336.

Dunstan, P.K., Foster, S.D. & Darnell, R. (2011) Model based grouping of species accross environmental gradient. *Ecological Modelling*, **222**(2011), 955–963.

Favrichon, V. (1994) Classification des espèces arborées en groupes fonctionnels en vue de la réalisation d'un modèle de dynamique de peuplement en forêt guyanaise. *Revue d'Écologie (Terre et Vie)*, **49**, 379–403.

Favrichon, V. (1998) Modeling the dynamics and species composition of tropical mixed-species uneven-aged natural forest: effects of alternative cutting regimes. *Forest Science*, **44**(1), 113–124.

Fieberg, J. & Ellner, S.P. (2001) Stochastic matrix models for conservation and management: a comparative review of methods. *Ecology Letters*, **4**(3), 244–266.

Gilks, W.R., Richardson, S. & Spiegelhalter, D.J. (eds.) (1996) *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London.

Gitay, H. & Noble, I.R. (1997) What are functional types and how should we seek them? *Plant Functional Types: Their Relevance to Ecosystem Properties and Global Change* (eds. T.M. Smith, H.H. Shugart & F.I. Woodward). No. 1 in *International Geosphere-Biosphere Programme*, Cambridge University Press, Cambridge, UK, pp. 3–19.

Gourlet-Fleury, S., Cornu, G., Jésel, S., Dessard, H., Jourget, J.G., Blanc, L. & Picard, N. (2005) Using models for predicting recovery and assessing tree species vulnerability in logged tropical forests: A case study from French Guiana. *Forest Ecology and Management*, **209**(1-2), 69–85.

Gourlet-Fleury, S., Guehl, J.M. & Laroussinie, O. (eds.) (2004) *Ecology and Management of a Neotropical Rainforest. Lessons Drawn from Paracou, a Long-Term Experimental Research Site in French Guiana*, Elsevier, Paris.

Gourlet-Fleury, S. & Houllier, F. (2000) Modelling diameter increment in a lowland evergreen rain forest in French Guiana. *Forest Ecology and Management*, **131**(1-3), 269–289.

Guillot, G., Estoup, A., Mortier, F. & Cosson, J. (2005) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.

Hauser, C., Cooch, E. & Lebreton, J. (2006) Control of structured populations by harvest. *Ecological Modelling*, **196**, 462–470.

Hooten, M., Wikle, C., Dorazio, R. & Royle, J. (2007) Hierarchical spatiotemporal matrix models for characterizing invasions. *Biometrics*, **63**, 558–567.

Jeffreys, H. (1946) An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London, Series A*, **186**(1007), 453–461.

Lieberman, D., Lieberman, M., Hartshorn, G. & Peralta, R. (1985) Growth rates and age-size relationships of tropical wet forest trees in costa rica. *Journal of Tropical Ecology*, **1**(2), 97–109.

Mao, C., Colwell, R. & Chang, J. (2005) Estimating the species accumulation curve using mixtures. *Biometrics*, **61**(2), 433–441.

Marin, J.M., Mengersen, K. & Robert, C.P. (2005) Bayesian modelling and inference on mixtures of distributions. *Bayesian Thinking, Modeling and Computation* (eds. D. Dey & C.R. Rao). No. 25 in *Handbook of Statistics*, Elsevier, Amsterdam, The Netherlands, pp. 459–507.

McLachlan, G.J. & Krishnan, T. (2008) *The EM Algorithm and Extensions*, 2nd edn. Wiley Series in Probability and Statistics, John Wiley & Sons, Hoboken, NJ.

McLachlan, G. & Peel, D. (2000) *Finite Mixture Models*. Wiley Series in Probability and Statistics, Wiley-Interscience, New York, NY.

Meilă, M. (2007) Comparing clusterings - an information based distance. *Journal of Multivariate Analysis*, **98**(5), 873–895.

549 Nascimento, H., Laurance, W., Condit, R., Laurance, S., D'Angelo, S. & Andrade, A. (2005)
550 Demographic and life-history correlates for amazonian trees. *Journal of Vegetation Sci-*
551 *ence*, **16**, 625–634.

552 Nobile, A. (2005) Bayesian finite mixtures: a note on prior specification and posterior com-
553 putation. *Technical Report 05-3*, University of Glasgow.

554 Phillips, P., Yasman, I., Brash, T. & van Gardingen, P. (2002) Grouping tree species for anal-
555 ysis of forest data in Kalimantan (Indonesian Borneo). *Forest Ecology and Management*,
556 **157**, 205–216.

557 Picard, N., Köhler, P., Mortier, F. & Gourlet-Fleury, S. (2012) A comparison of five classifi-
558 cations of species into functional groups in tropical forests of French Guiana. *Ecological*
559 *Complexity*, in press. Http://dx.doi.org/10.1016/j.ecocom.2012.03.003.

560 Picard, N., Mortier, F., Rossi, V. & Gourlet-Fleury, S. (2010) Clustering species using a model
561 of population dynamics and aggregation theory. *Ecological Modelling*, **221**(2), 152–160.

562 Pledger, S., Pollock, K.H. & Norris, J.L. (2010) Open capture-recapture models with hetero-
563 geneity. II: Jolly-Seber model. *Biometrics*, **66**(3), 883–890.

564 Poorter, L., Wright, S.J., Paz, H., Ackerly, D.D., Condit, R., Ibarra-Manríquez, G., Harms, K.
565 E. and. Licona, J.C., Martínez-Ramos, M., Mazer, S.J., Muller-Landau, H.C., Peña-Claros,
566 M., Webb, C.O. & Wright, I.J. (2008) Are functional traits good predictors of demographic
567 rates? Evidence from five neotropical forests. *Ecology*, **89**, 1908–1920.

568 Pritchard, J., Stephens, M. & Donnelly, P. (2000) Inference of population structure using
569 multilocus genotype data. *Genetics*, **155**, 945–959.

570 R Core Team (2012) *R: A Language and Environment for Statistical Computing*. R Founda-
571 tion for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org`.
572 ISBN 3-900051-07-0.

573 Richardson, S. & Green, P.J. (1997) On Bayesian analysis of mixtures with an unknown
574 number of components. *Journal of the Royal Statistical Society, Series B*, **59**(4), 731–792.

575 Robert, C.P. & Casella, G. (2005) *Monte Carlo Statistical Methods*. Springer Texts in Statis-
576 tics, Springer, New York, NY.

577 Sebert-Cuvillier, E., Paccaut, F., Chabrerie, O., Endels, P., Goubet, O. & Decocq, G. (2007)
578 Local population dynamics of an invasive tree species with a complex life-history cycle: A
579 stochastic matrix model. *Ecological Modelling*, **201**, 127–143.

580 Steneck, R. & Dethier, M. (1994) A functional-group approach to the structure of algaldom-
581 inated communities. *Oikos*, **69**, 476–498.

582 Stott, I., Townley, S., Carslake, D. & Hodgson, D.J. (2010) On reducibility and ergodicity of
583 population projection matrix models. *Methods in Ecology and Evolution*, **1**(3), 242–252.

584 Swaine, M. & Whitmore, T. (1988) On the definition of ecological species groups in tropical
585 rain forests. *Vegetatio*, **75**, 81–86.

586 Turner, I.M. (2001) *The Ecology of Trees in the Tropical Rain Forest*. Cambridge Tropical
587 Biology Series, Cambridge University Press, Cambridge, UK.

588 Usher, M.B. (1966) A matrix approach to the management of renewable resources, with
589 special reference to the selection forests. *Journal of Applied Ecology*, **3**, 355–367.

590 Usher, M.B. (1969) A matrix model for forest management. *Biometrics*, **25**(2), 309–315.

591 Supporting Information

592 Additional Supporting Information may be found in the online version of this article.

593 **Data S1.** R scripts for the Bayesian inference algorithm.

594 **Data S2.** Parameters of the mixture matrix models with five tree species groups at Paracou,
595 French Guiana.

Table 1: Hyper-prior distributions of the parameters used for simulations. $\mathcal{D}$ is the Dirichlet distribution, $\mathcal{G}$ is the gamma distribution. 'Var' is the variance of d_i, of p_l, q_l, m_l, and of f, respectively.

Differentation	Diameter $\vec{d}$		Transition (p_l, q_l, m_l)		Fecundity f	
Level	Distribution	Var	Distribution	Var	Distribution	Var
Ldiff_1	$\mathcal{D}(1,1,1,1,1,1,1,1)$	0.0121	$\mathcal{D}(1,1,1)$	0.055	$\mathcal{G}(10,1000)$	10^{-5}
Ldiff_2	$\mathcal{D}(3,3,3,3,3,3,3,3)$	0.0044	$\mathcal{D}(3,3,3)$	0.022	$\mathcal{G}(10,2000)$	$2.5{\cdot}10^{-6}$
Ldiff_3	$\mathcal{D}(5,5,5,5,5,5,5,5)$	0.0027	$\mathcal{D}(5,5,5)$	0.014	$\mathcal{G}(10,3000)$	$1.1{\cdot}10^{-6}$
Ldiff_4	$\mathcal{D}(7,7,7,7,7,7,7,7)$	0.0019	$\mathcal{D}(7,7,7)$	0.010	$\mathcal{G}(10,4000)$	$6.25{\cdot}10^{-7}$
Ldiff_5	$\mathcal{D}(9,9,9,9,9,9,9,9)$	0.0015	$\mathcal{D}(9,9,9)$	0.008	$\mathcal{G}(10,5000)$	$4{\cdot}10^{-7}$

Table 2: Comparison between simulated and estimated classifications: mean of (I_1, I_2) on the 50 simulations for 100 individuals per species, depending of the differentiation levels for the hyper-priors. Definition of the Ldiff_i is given in Table 1. n.d. means "not defined".

Differentiation level	1 group	5 groups	10 groups
Ldiff_1	(1,1)	(1,1)	(1,1)
Ldiff_2	n.d.	(0.996,0.996)	(0.998,0.988)
Ldiff_3	n.d.	(0.996,0.989)	(0.978,0.889)
Ldiff_4	n.d.	(0.983,0.933)	(0.929,0.686)
Ldiff_5	n.d.	(0.964,0.865)	(0.899,0.574)

Table 3: Observed vital rates of groups (Obs.) and average vital rates computed from the estimated transition rates (Est.): two-year dbh increment ($\Delta\mathrm{DBH}$), two-year mortality rate, two-year fecundity rate, upper bound of diameters ($\mathrm{DBH95}$) and two-year turnover of the five groups obtained using matrix population mixture model classification. The observed $\Delta\mathrm{DBH}$ for group i was $\frac{1}{k_i}\sum_{j=1}^{k_i}(Y_j^{1995} - Y_j^{1993})$, where Y_j^t was the dbh of individual j at year t, and k_i the number of individuals in group i.

	ΔDBH (cm)		Mortality (%)		Fecundity (%)		DBH95 (cm)		Turnover (%)	
Group	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
1	0.38	0.42	0.91	1.31	1.25	1.25	65.3	68.1	1.08	1.28
2	0.27	0.25	1.33	1.58	1.04	1.05	44.2	45.6	1.19	1.32
3	0.24	0.24	2.34	2.70	1.02	1.09	37.4	37.8	1.68	1.90
4	0.13	0.10	2.21	2.38	1.54	1.47	24.2	24.7	1.87	1.93
5	0.08	0.05	2.18	2.74	1.86	2.03	16.4	17.9	2.02	2.39

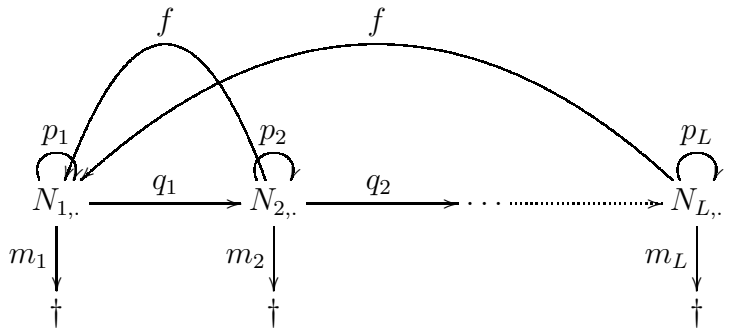

Figure 1: Life cycle representation of the Usher projection matrix model, where p_l is the probability for an individual to stay in class l, q_l is the probability to move up from class l to $l+1$, m_l is the probability of dying and f is the average fecundity.

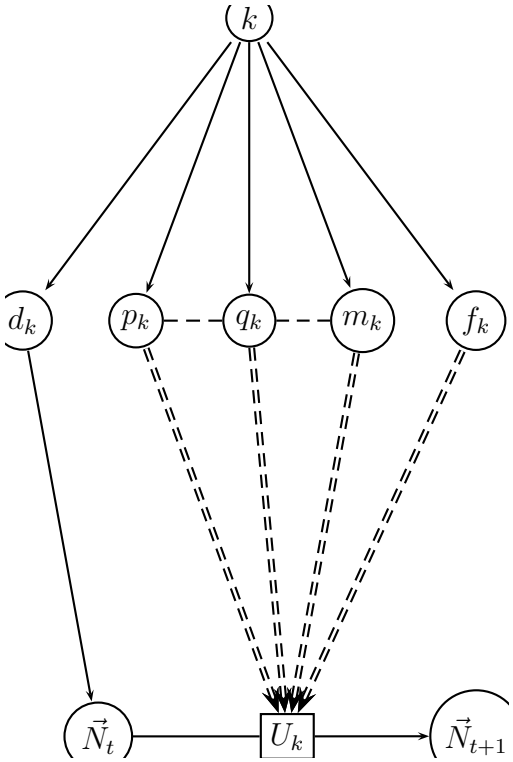

Figure 2: Direct acyclic graph of the mixture of Usher projection matrix model. Double dot arrows indicate deterministic links, dot lines indicate direct links, circles indicate random nodes and frames indicate deterministic nodes.

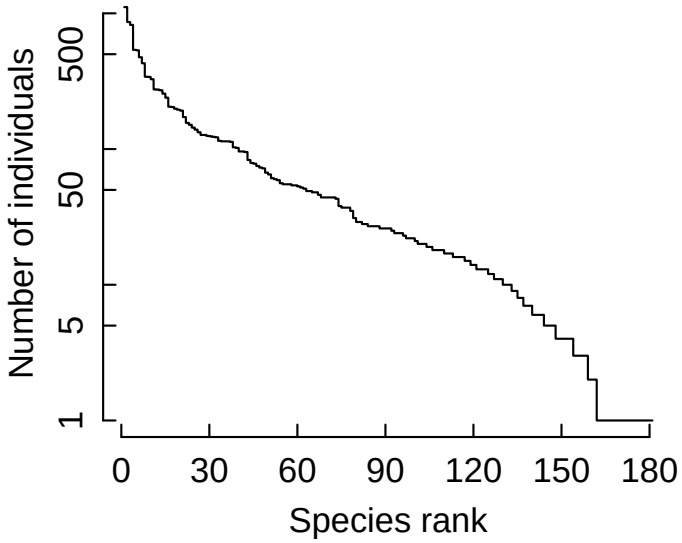

Figure 3: Rank-abundance diagram in the control plots at Paracou in 1993.

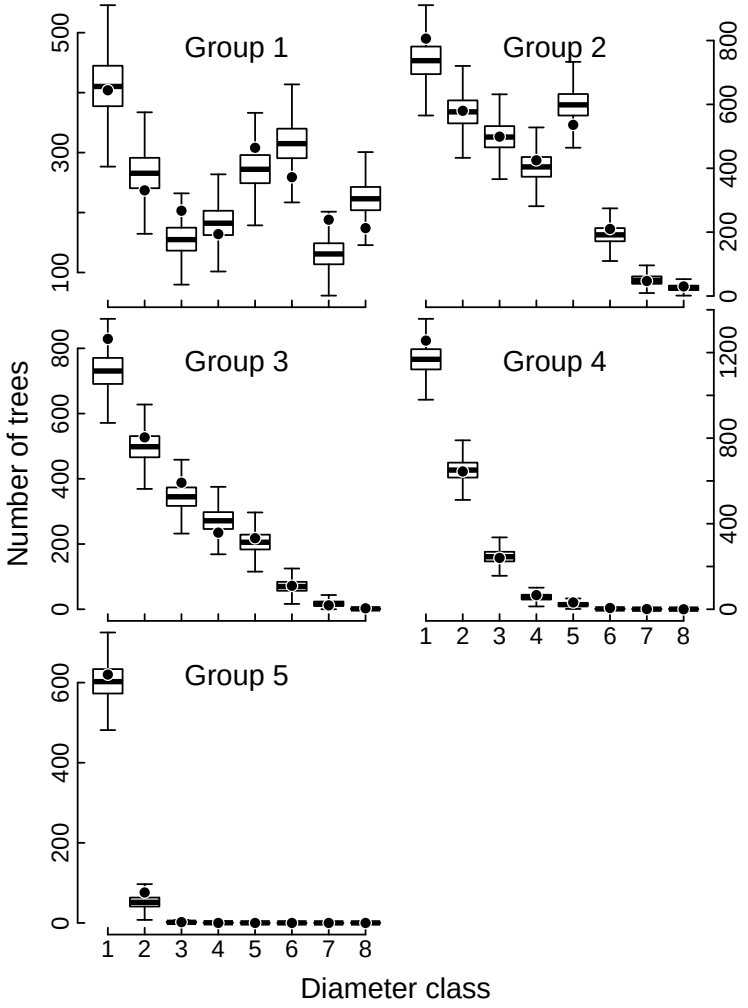

Figure 4: Predicted (boxplot) and observed (black dot) number of individuals in each diameter class and each species group in the control plots at Paracou in 2009.

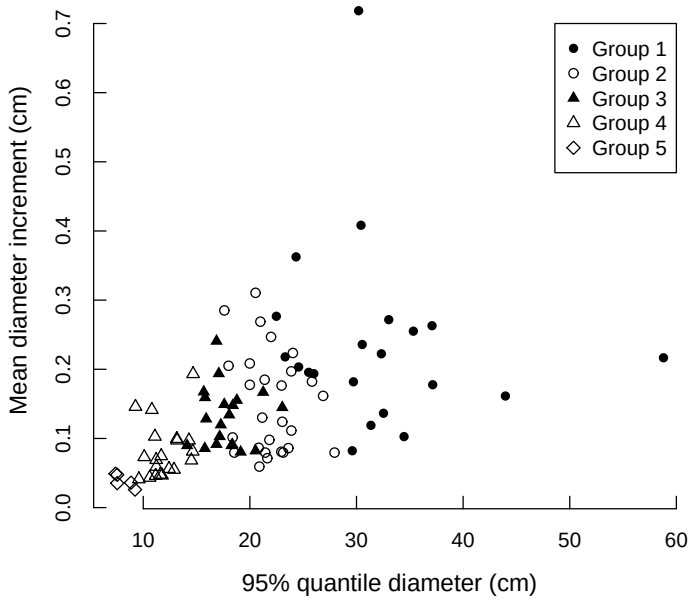

Figure 5: Upper bound of diameters (95% quantile of dbh in 1995, in cm) versus mean diameter increment between 1993 and 1995 (cm) for 93 species at Paracou, French Guiana. The five different symbols correspond to the five groups defined by the mixture matrix model.